DOCUMENT RESUME

ED 254 953 EA 017 619

AUTHOR Grobe, Robert P.; And Others
TITLE Developing a School Improvement Plan: Integrating a Large Assessment System with Instruction.
PUB DATE Apr 84
NOTE 45p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 23-27, 1984).
PUB TYPE Reports - Descriptive (141) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Accountability; Achievement Tests; Criterion Referenced Tests; *Curriculum Design; Curriculum Development; *Educational Testing; Elementary Secondary Education; *Instructional Development; Minimum Competency Testing; Models; Norm Referenced Tests; School Based Management; Standardized Tests; *Student Evaluation
IDENTIFIERS Dallas Independent School District TX

ABSTRACT

The Dallas (Texas) Independent School District has developed an assessment system that is fully integrated into its learning system, which also includes planning, instruction, and accountability. This paper describes the learning system in Dallas in terms of a four-part model: planning, instruction, assessment, and accountability. Following this, the three governing principles of the learning system are delineated: (1) as much control as possible should reside at the building level; (2) flexibility and change must be allowed; and (3) activities asssociated with the learning model should be as unobtrusive as possible. The third section discusses the minutiae, trivia, and general headaches that must be overcome to implement this program successfully. Guidelines are accordingly provided for item formatting and test assembly. It is suggested that this learning syseem works in the Dallas Independent School District because (1) assessment is part of the system, (2) the system is governed by the three principles identified above, and (3) some measure of luck has prevailed. An appendix details test development activities for survey and classroom tests. (TE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED254953

Developing A School Improvement Plan: Integrating a Large Assessment System with Instruction

Robert P. Grobe, Jerrilyn V. Andrews, Deberie G. Adkins, & Dean G. Arrasmith

Dallas Independent School District

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Robert P. Grobe
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Persons associated with administering and implementing assessment programs frequently hear comments similar to the following:

1. "Testing interrupts instruction."
2. "Testing takes too much time."
3. "Results from our testing program make schools and teachers look bad."
4. "The tests don't provide any practical information."
5. "The tests do not test what we teach."

As suggested in these comments, there are problems in getting data from assessment programs used.

The Dallas Independent School District (DISD) was not immune to the problems surrounding the use of assessment data. As was the case in many large-city districts, Dallas has witnessed a reduction in the size of its Department of Research and Evaluation as a direct result of the failure of administrators (both central and local building) and teachers to use data made available from the assessment programs.

However, the DISD is progressing from the position of having an assessment program that does not work to the position of having an assessment program that <u>does</u> work. The concept (or perception) of

---

Paper presented at the annual meeting of the American Educational Research Association in New Orleans, Louisiana, April 1984.

EA 017 619

assessment, as an entity that is separate and distinct from instruction, has changed to a view of assessment as part of a continuum of instruction. In other words, assessment has become an integral part of a learning system which also includes planning, instruction, and accountability. Some of the more overt changes which have occurred during the process of integrating assessment into a continuum of instruction include: (a) the measurement instruments which make up the testing program have changed; (b) the administrative structure into which assessment data are fed has been reorganized; (c) the relationship between the Departments of Research and Evaluation and Curriculum and Instruction has substantially changed; and (d) the state, the community, and the Board of Education have increased expectations and demands relative to improving student achievement. The changes in all four of these elements were necessary conditions before a complete learning system could begin to operate in the DISD.

The purpose of this discussion is to describe (a) the learning system in terms of a four-part model, (b) three governing principles which have been used to implement activities comprising the learning system and (c) the minutia, trivia, and general headaches which must be worked through and overcome to have successful implementation. The four parts of the model are planning, instruction, assessment, and accountability. The governing principles are: (a) as much control as possible should reside with personnel at the local building, (b) flexibility and change must be allowed, and (c) limited obtrusiveness is necessary for ease of implementation. It is believed that this learning system works in DISD because: (a) assessment is part of the

system, (b) the system is governed by the three principles identified above, and (c) some measure of luck has prevailed.

THE MODEL

The learning system model which is to be described in the following paragraphs consists of four parts: (a) planning, (b) instruction, (c) assessment, and (d) accountability. Each of the parts will be explained briefly such that the relationship or integration can be recognized.

Planning

The first part of the model is planning. In DISD, annual planning occurs at the local building level. Each campus develops a school improvement plan which focuses on improving student achievement in areas designated as priorities on the basis of the system-wide assessment program. The planning cycle begins during the final two months of the instructional year as the different types of assessment data become available in the schools. While these data are being used to evaluate the previous year's instructional strategies, they are also used to determine the needs for instructional focus during the next school year.

The results of the state criterion-referenced tests and the local criterion-referenced tests are the most useful information the schools have during this early plan development/revision stage. These data specifically illustrate to the faculty which skills continue to show deficiencies and thus which skills must be the foci of their plan. Since the state tests are administered only at grades 3, 5, and 9, the local

tests are the source of the information necessary to plan for instruction by grade levels and across grade levels.

Included in the school improvement plan, there must be a description of the means by which the school administrators will evaluate the instructional strategies on an interim basis. This may be written into the instructional strategies themselves, or it may be written as an element of the internal monitoring system for the plan. The interim evaluation is, in most cases, the use of the District's criterion-referenced classroom tests.

Instruction

A six-step instructional process is used in all classes. The strategies in the school improvement plan provide for additional instruction in priority areas where skills are the weakest. These instructional strategies must be written in a six-step teaching format which includes: (a) find out what students already know, (b) tell the students what the objective is, (c) demonstrate/illustrate/show the students the skill, (d) guide the students as they practice the skill, (e) allow students to apply the skill independently, and (f) evaluate mastery level. Again, the progress of students in the classroom and the first and last steps of the six-step process (find and evaluate) require the presence of appropriate assessment instruments which are available in the classroom and consistent across the District.

Assessment

The assessment system in the DISD includes both norm-referenced and criterion-referenced tests. Considerable effort has been invested to make the assessments more user-oriented (teachers and administrators).

Test answer sheets are pre-marked with name and 'demographic' information, where possible; test scanning and scoring is provided within the District; and student, class, school, and District test score reports are routinely provided, while other reports (e.g., special groups, skills analysis) are provided upon request.

Norm-referenced tests. The DISD relies on the Iowa Tests of Basic Skills (ITBS) and Tests of Achievement and Proficiency (TAP) for norm-referenced measures of student progress in grades K-11. Certain of the subtests are required by the District while the remaining subtests are given at the option of the individual schools. Routinely, student, class, school, and District score reports are generated immediately following the spring testing. The data are then reaggregated and reported by the next year's fall enrollment.

Criterion-referenced survey tests. In conjunction with the norm-referenced tests, criterion-referenced tests are given in grades 1-6. These tests measure attainment of specific mathematics and communication curriculum objectives that have been identified as promotion standards. Reports are generated for teachers, principals, subdistrict administrators, and district administrators. The data reported include: (a) the performance of each student, (b) class summaries, (c) school summaries, (d) subdistrict summaries, and (e) a district summary. These district- wide survey tests with the attendant reporting system provide the standardized student performance information on the promotion standards that is needed by classroom teachers and by school, subdistrict, and district administrators for the accountability and planning parts of the model.

Appendix A contains a detailed description of the survey tests and their characteristics. In addition the development process is discussed.

Criterion-referenced classroom tests. Criterion-referenced tests that measure mathematics and communications curriculum objectives are provided to the teachers. These tests are designed for instructional use in the classroom. They provide standard measures of student progress in a continuum of instruction. Teachers can use any part or parts of the tests in the order that best fits the order of instruction. Teachers are often particularly interested in using those parts of the classroom tests that correspond to end-of-year accountability measures (i.e., criterion-referenced survey tests and state criterion-referenced tests).

A scoring and reporting service is provided that relies on a central computer for scanning, scoring, and reporting test scores. Student performance as measured by the tests can be continually updated and stored over the school year. Teachers can test their students as frequently as they need, covering only the objectives they need to measure for attainment. These classroom tests are of considerable assistance to some teachers as they are required by DISD policy to maintain a record of student progress by curriculum objectives.

Appendix A contains a detailed description of the classroom tests. Different applications for the classroom tests are also discussed.

Accountability

This section is concerned with describing the accountability part of the learning system model. The discussion will center on (a) the definition of accountability as used in DISD and (b) activities in the process of accountability.

Definition of accountability. A prerequisite to the discussion of the accountability part of the learning system is an operational definition of accountability as it is used in the DISD. To be accountable in this process is to be responsible. There are three basic steps which, when carried out, demonstrate that responsibilities have been successfully discharged. Step one consists of stating the objectives of instruction in an explicit and measureable way. The second step is to state, in specific terms, the strategies which will be employed to attain the objectives. Step three is the process of documenting the implementation of the strategies. Having followed the three basic steps, accountability is demonstrated by the existence of one of the following conditions:

1. The objectives are attained.

OR

2. The objectives are not attained, but the documentation verifies that the strategies were implemented as planned.

OR

3. The objectives are not attained and there were extenuating circumstances that prevented either the implementation of strategies as planned or the result of that implementation from being manifested. Examples of such circumstances might be: drastic population change, loss of funds and/or personnel, absenteeism, failure of materials and resources to arrive, etc.

The accountability process. Schools must establish measurable student achievement objectives for each goal contained in the school improvement plan. These goals are in the areas of mathematics, reading, and writing/language arts. While certain objectives are required to be based on state criterion-referenced test and the norm-referenced tests, schools may also use objectives based on the locally-developed,

criterion-referenced tests. The instructional strategies are then designed to bring about the improvement in student achievement necessary to attain the stated objectives.

Accountability is required at several levels of the DISD structure. Not only are teachers and school building administrators held accountable for the objectives in the school improvement plans, but also the subdistrict administrators and other central staff including individuals from the Department of Research and Evaluation (R&E), and Curriculum and Instruction (C&I) bear a portion of the responsibility. After the teachers and school administrators have determined what they feel their objectives should be, a conference is held with the subdistrict administrators, technical personnel from R&E and instructional personnel from C&I. The objectives must be agreed upon by the entire group. Likewise, the instructional strategies must receive the approval of the subdistrict and central team.

Recall that the school includes in its plan the means for internal monitoring and interim evaluation. This allows the school to insure itself that the strategies are being implemented. As stated earlier, the classroom tests serve as immediate measures of progress resulting from the implementation of the instructional strategies in the priority skill areas.

During the course of the school year, the subdistrict and central staff teams monitor the implementation of the school improvement plans by making two rounds of classroom visitations. A classroom observation form is completed for every classroom visited. These are compiled by each monitor into a building report which is then summarized in a team

report. Documentation and concrete evidence must be available to show that the school is implementing the plan.

All of the available criterion-referenced measures and the norm-referenced test results serve as the ultimate mechanism for determining whether or not the first condition of accountability exists: The objectives are attained. The results of the monitoring visits can be reviewed to determine whether the second or third conditions of accountability pertain. As previously indicated accountability flows directly back into the planning for the succeeding school year thus completing a cycle.

THE PRINCIPLES OF IMPLEMENTATION

This section of the paper is concerned with discussing the principles which govern the implementation of the learning model. Three general principles are described: (a) as much control as possible should reside with personnel at the local building, (b) change and flexibility must be allowed, and (c) limited obtrusiveness is necessary for ease of implementation.

Local Building Control

Simply stated the principle of local building control calls for allowing the personnel at each school to have as much autonomy as possible. Activities related to the phases of the model are used as examples of local control in the following paragraphs.

Instructional decentralization. The personnel at each school have certain District emphases that they must address; however, they have a large amount of control in addressing local priorities and in how they

they address the District emphases. They have control over allocating their budgeted dollars among their priority needs. They also have control over the kind of staff development they will have. In addition, they can design their own monitoring system related to their goals.

Planning and accountability. The description of planning and accountability in the model illustrates the importance of local control. While the District recommends the priority areas, schools are able to develop their school improvement plans based on individualized achievement data so that the efforts are directed to the appropriate areas where weaknesses in skills exist. For instance, if an individual school finds that one of the District's weak reading skills is not a problem for the students in that building but that some other reading skill is a problem, they are free to focus on that weak skill. The overall goal area is reading but the individual school's instructional foci may vary.

School personnel have come to recognize the importance of criterion-referenced measures in developing a school plan. Because the system-wide test results have a lag time in getting back to the buildings in the spring when the plans must be developed/revised many schools depend on the optional local criterion-referenced survey test results.

Assessment and evaluation. Each school exerts a measure of local control in designing their assessment system. The schools have some indirect control over the way they use the norm-referenced test data in that they can select a variety of report formats. The schools have absolute control over the use of the classroom tests. Personnel in each school may choose whether or not to use classroom tests. If they choose

to use them, they make all decisions concerning how and when they are administered and how they will use the results.

Flexibility and Change

Because of federal and state mandates, technological change, and societal and political pressures, education must allow for flexibility and change. The parts of the instructional model are discussed in terms of how they allow for flexibility and change.

Curriculum/instructional changes. Typically, changes have been made in the DISD curriculum to match (a) mandates of the Texas Education Agency, (b) assessment systems, and (c) new trends in instruction. The major problem related to changes in curriculum and instruction has not been related to curriculum, but to implementing corresponding changes in the instruction that takes place in the schools. The planning and accountability phases of our model are now used as vehicles to force changes in the instruction. More specifically, monitoring teams that visit the schools watch for the necessary changes in instruction and, if they are not found, the building principal is held responsible.

Assessment flexibility and changes. By having test development personnel within the district, changes in the assessment system can be incorporated relatively quickly to match changes in curriculum and instruction. For instance, if objectives are inserted to match state mandates, then tests can be revised in an ongoing manner. Also, these individuals can respond to special requests for reaggregating data.

In scoring classroom tests for teachers, efforts are made to keep the service as flexible as possible. Teachers are allowed to test as few

as one objective or as many as all objectives on the test. They may also test as many times during the year as they like.

Flexibility in school improvement planning and accountability. While the DISD commits to a set of goals for a five-year period of time, the annual objectives may change from year to year. For example, the state criterion-referenced tests may indicate that, in the area of reading, the skills which require the focus of the District in a given year are: (a) identifying the main idea, (b) sequencing events, and (c) making generalizations. Because of improvement in one of these skills, the succeeding year may include a and b but c would be replaced by "distinguishing fact from nonfact." It may be that a fourth focus is added. The magnitude of the objectives may also change from year to year. If a school makes large gains in achievement in a skill area one year, they may set a smaller objective for that skill the following year.

Limited Obtrusiveness

In limiting obtrusiveness, efforts were made to make activities associated with the learning model as unnoticeable, unshowy, and inconspicuous as possible. Some of the ways this was done are discussed as examples below.

Instruction and curriculum. Since teachers often resist change, and since changes in instructional methods and curriculum directly impact the lives and teaching styles of teachers, such changes must be made in a way that is relatively inconspicuous. Admittedly, this cannot always be done. However, two methods have been used. First, curriculum changes were made to fit formats used in District-adopted texts whenever possible. Second, and perhaps more importantly, the school improvement

planning process was used to introduce exemplary instructional practices to teachers. Because the school improvement plans are locally developed, school personnel tend to have a vested interest in the planned instruction and objectives. In this way, instructional improvement can occur without the appearance of a massive change.

Planning and accountability. Planning and accountability activities are basically obtrusive by nature; however, efforts to limit or minimize the obtrusiveness include (a) changing the process as little as possible from year to year and (b) providing as much assistance as possible.

As indicated earlier, the DISD commits to priority goals for a period of at least five years. The priority goals may be, for example, to improve student achievement in reading, writing/language arts, and mathematics. Thus, the annual school improvement plans consistently focus upon these goals. Additionally, the format of the school improvement plans and the requirements in terms of description of instructional processes change very little from year to year thereby allowing schools to revise and refine rather than constantly redevelop plans from a beginning point.

The school improvement planning process and the external monitoring process are designed with the intent of providing assistance to the schools, not with the intent of obstructing progress through unwarranted intrusion. It is assumed that the local building personnel do have the best idea of what the instructional program should be in order to improve student performance.

Assessment. In attempting to limit obtrusiveness of the assessment system, three efforts have been made. District personnel responsible for assessment have tried to (a) foster a view of assessment as an integral

part of the learning system, (b) fit existing needs for information, and (c) be service-oriented.

To foster a view of assessment as a part of a continuum, efforts have been made to show local school personnel how test information can be used to focus instruction on priority needs. In addition, school personnel have been educated on the use of classroom tests in conducting interim assessments. That is, ongoing, short term testing conducted after each objective and/or groups of objectives.

To fit existing needs for information and to be service-oriented, efforts have been made to computer-generate reports that reduce teacher paperwork and to present reports in formats with which schools were familiar or which had been requested by schools. For example, the classroom tests are available for teachers in two basic regards: (a) Student Progress Reports and (b) Class Profiles. Samples of these computer-generated reports for Level 6 mathematics and communications are presented in Figures 1-4. When needed, student progress reports and class profiles are available from the computer data base. This (optional) service results in a considerable reduction of classroom record-keeping (recall what the District requires records of student progress). District policy requires teachers to have a minimum of two parent conferences each year. Student Progress Reports, used in parent conferences, are computer generated for each teacher using the computer scoring service.

GRADE 6

Mathematics Baseline Objectives

dallas independent school district

STUDENT PROGRESS REPORT

MARKING CODE	
[X]	MASTERED (3,4)
[/]	NOT MASTERED (0,1,2)
[ ]	NOT INTRODUCED OR NOT TESTED

BAIL JAYME (Student) SHELT VIC (Teacher) 6/2 (Grade/Section) DAVID G. BURNET (School) 12/10/82 (Date Scored)

Placement Survey: Pretest Date _____ Grade Equivalent ___ Posttest Date _____ Grade Equivalent ___

DISD Promotion/Retention (P/R) Standards are printed in "ALL CAPS." See Mathematics Baseline for a detailed explanation of the objectives.

A. NUMERATION
- [X] 1. Interpret place value
- [X] 2. Read & write numbers & word names
- [/] 3. Write power of 10 in exponential form
- [X] 4. Write numbers in standard & expanded form
- [X] 5. Order & compare whole numbers
- [X] 6. Round to the nearest 10, 100, 1,000, 10,000, 100,000
- [X] 7. Interpret place value of decimals to 10,000ths
- [/] 8. Compare decimals to 10,000ths
- [X] 9. Apply concepts of rounding through 1,000ths in solving problems
- [X] 10. INTERPRET PLACE & TOTAL VALUE IN 5-DIGIT NUMBER WITH UP TO 3 DECIMAL PLACES

B. ADDITION & SUBTRACTION
- [/] 11. ADD WHOLE NUMBERS (4 DIGITS OR LESS)
- [X] 12. Solve open sentences (+,-)
- [X] 13. SUBTRACT WHOLE NUMBERS (4 DIGITS OR LESS)
- [X] 14. ADD & SUBTRACT DECIMALS (TO 1,000THS)
- [X] 15. Solve money & word problems (+,-)

C. STRUCTURE AND MULTIPLICATION
- [X] 16. Use commutative, associative & distributive properties (+,x)
- [X] 17. MULTIPLY WHOLE NUMBERS (TO 999X99)
- [/] 18. Multiply decimals to 1,000ths

D. MEASUREMENT AND GEOMETRY I
- [X] 19. SELECT APPROPRIATE UNIT OF CUSTOMARY MEASURE
- [X] 20. SELECT APPROPRIATE UNIT OF METRIC MEASURE
- [X] 21. FIND PERIMETER OF GEOMETRIC FIGURES
- [/] 22. Find area, volume, & surface area of rectangular geometric figures

E. DIVISION
- [X] 23. Solve open sentences (x,/)
- [X] 24. DIVIDE WHOLE NUMBERS (TO 999/99)
- [X] 25. Divide decimals to 100ths

F. PROBLEM SOLVING I
- [X] 26. Complete open sentences by using symbol or number
- [X] 27. Write number sentence using parentheses & order of operations
- [/] 28. Estimate products & quotients
- [X] 29. Solve money & word problems (x,/)
- [X] 30. Classify numbers as prime/composite & determine factors (to 99)

F. PROBLEM SOLVING I (continued)
- [X] 31. Identify factors & greatest common factor of set of numbers
- [/] 32. Locate points & coordinates of points on a number line

G. FRACTIONS AND MIXED NUMBERS
- [X] 33. RENAME A FRACTION IN SIMPLEST FORM THROUGH 20THS
- [/] 34. Order & compare fractions & mixed numbers
- [X] 35. Find fractions equivalent to given fraction
- [/] 36. Add & subtract fractions & mixed numbers with like & unlike denominators

H. MEASUREMENT AND GEOMETRY II
- [X] 37. Measure & classify angles & triangles by degree measure
- [X] 38. Describe parallelograms & triangles
- [X] 39. Identify parts of a circle

I. FRACTIONS
- [X] 40. MULTIPLY FRACTIONS WITH NUMERATORS & DENOMINATORS LESS THAN 10 & SIMPLIFY
- [/] 41. Divide & multiply fractions/whole numbers by fractions

J. PROBLEM SOLVING II
- [/] 42. SOLVE WORD PROBLEMS (+,-,x,/) WITH WHOLE NUMBERS; (+,-) WITH DECIMALS
- [/] 43. Find average of numbers (to 5,999)
- [X] 44. READ & INTERPRET PICTURE, LINE & BAR GRAPHS & TABLES

K. MEASUREMENT
- [/] 45. Compute & simplify customary & metric measures

L. RATIO AND PERCENT
- [/] 46. Express fractions as decimals or percents & vice versa

M. GEOMETRY
- [X] 47. Compute areas of triangles, squares, rectangles, parallelograms, & trapezoids
- [/] 48. Determine circumference & area of circle

N. PROBLEM SOLVING III
- [X] 49. Solve money problems including cost comparison

O. NUMERATION
- [/] 50. Compare other numeration systems to the decimal system

Fall Conference Date: _____ _____ Parent/Guardian

Spring Conference Date: _____ _____ Parent/Guardian

Figure 1. Sample student progress report for mathematics.

BEST COPY

BEST COPY

CLASS PROFILE

MATHEMATICS GRADE 6

School: DAVID O. BURNET

Teacher: SHELT VIC

STUDENT NAMES		1 Place value	2 Number & word names	3 Powers of 10	4 Standard & expanded form	5 Order & compare numbers	6 Round numbers	7 Decimal place value	8 Compare decimals	9 Round decimals	10 PLACE VALUE (6 DIGITS, 3 DECIMAL PLACES)	11 ADD WHOLE NUMBERS	12 Open sentences (+,-)	13 SUBTRACT WHOLE NUMBERS	14 DECIMAL PROBLEMS (+, -)	15 Story problems (+,-)	16 Commutative, associative & distributive properties	17 MULTIPLICATION (TO 999X99)	18 Multiply decimals	19 CUSTOMARY MEASURE	20 METRIC MEASURE	21 FIND PERIMETER	22 Find area, volume, & surface area	23 Open sentences (x, /)	24 DIVISION (TO 999/99)	25 Division (to 100ths)	26 Open sentences (symbol or number)	27 Open sentences using () & order of operations	28 Estimation (x, /)	29 Story problems (x, /)	30 Prime/composite numbers	31 Identify factors & greatest common factor	32 Points on number line	33 RENAME FRACTION	34 Order/compare fractions & mixed numbers
BAIL	JAYME	X	X	/	X	X	X	X	/	X	X	/	X	X	X	X	X	X	/	X	X	X	/	X	X	X	X	/	X	X	X	/	X	/
BELL	DEWAYNE	X	X	X	/	/	X	/	/	/	/	X	X	X	X	/	/	X	X	X	X	/	X	/	X	X	X	X	/	/	/	X	/	/
BOAT	KEVIN	X	X	/	/	X	/	/	/	X	X	/	X	X	X	/	/	/	/	X	X	/	/	/	/	/	/	/	/	/	/	X	X	/
BOUF	KONRAD	X	X	/	X	X	/	X	X	/	X	X	X	X	X	X	X	X	X	/	X	X	/	X	X	X	X	X	X	/	X	X	X	/
CHRIS	ALICE	X	X	/	/	/	X	/	/	/	X	X	X	X	X	X	X	X	X	X	X	X	/	X	X	X	X	/	X	/	X	/	/	/
CRON	VERONIC	X	X	X	X	X	X	X	/	X	X	X	X	X	X	X	X	X	X	X	X	X	/	X	X	X	X	X	X	/	X	X	X	/
DUVA	CHARLIE	X	X	/	X	X	X	X	/	/	X	X	X	X	X	X	X	/	X	X	X	/	X	/	X	X	X	/	X	X	X	/	X	/
EDDE	CHRIS	X	X	/	X	X	X	X	/	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	/
FERGUS	ROMEO	X	X	/	X	X	X	X	/	/	/	X	X	X	X	X	X	X	/	X	X	X	/	X	X	X	X	/	X	/	X	X	X	/
FOST	DEBBIE	X	X	/	/	/	/	X	/	X	X	X	X	X	X	/	/	X	X	X	X	/	/	X	X	/	X	X	X	/	/	/	/	/
FRODSH	MIKE	X	X	X	/	X	/	X	/	X	X	X	X	X	X	X	X	X	/	X	/	X	X	X	X	X	X	X	X	/	X	X	/	/
GRIMM	GANNON	X	X	X	X	X	/	X	/	X	X	X	X	X	X	X	/	X	/	X	X	X	X	X	X	X	X	X	/	/	/	/	/	/
HAGGER	KAREN	X	X	/	X	X	X	X	/	X	X	X	X	X	X	X	X	X	X	X	/	X	X	X	X	X	X	X	/	/	X	X	/	/
HILL	GERONAL	X	X	X	/	/	/	/	X	X	/	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	/	/	X	X	/	/
HILT	LORI	X	X	/	/	X	/	X	/	X	X	X	X	X	X	X	/	X	X	X	/	X	X	X	X	/	X	/	/	/	X	X	/	/
HUD	DARC	X	X	X	/	X	/	X	/	X	X	X	X	X	X	X	X	X	/	X	/	X	X	X	X	X	X	X	X	X	X	X	/	/
JONI	MARCUS	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	/	X	X	X	X	X	X	X	X	X	X	X	X	X
JON	AMIE	X	X	/	X	X	/	X	/	/	X	X	X	X	X	X	X	X	/	X	/	X	X	X	X	X	X	X	X	X	/	X	/	/
KEY	TROY	X	X	X	X	X	X	X	/	X	X	X	X	X	X	X	X	X	X	/	X	X	X	X	X	/	X	/	/	/	X	X	/	/
LEW	LACINDA	X	X	/	X	X	X	X	/	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	/	X	X	X	/	X	X	X	/
OVE	CATHY	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	/	X	X	X	X	X	X	/	/	/	X	X	X	/
PALAC	NANCY C	X	X	X	/	/	X	/	/	/	/	X	X	X	X	X	X	X	X	X	/	/	X	X	X	X	X	/	/	/	X	X	/	/
PEACO	MATTHEW	X	X	/	X	/	/	X	/	X	X	X	/	/	/	X	/	/	/	X	/	X	/	X	/	/	X	/	/	/	X	X	X	/
RIDING	SHERI	X	X	X	X	X	/	X	/	X	X	X	X	X	X	X	X	/	X	/	X	/	X	/	X	X	X	/	X	/	/	/	X	/

CONTINUED ON THE NEXT PAGE

Figure 2. Sample class profile for mathematics.

CLASS PROFILE

MATHEMATICS GRADE 6

School: DAVID O. BURNET

Teacher: SHELT VIC

STUDENT NAMES		1 Place value	2 Number & word names	3 Powers of 10	4 Standard & expanded form	5 Order & compare numbers	6 Round numbers	7 Decimal place value	8 Compare decimals	9 Round decimals	10 PLACE VALUE (6 DIGITS, 3 DECIMAL PLACES)	11 ADD WHOLE NUMBERS	12 Open sentences (+,-)	13 SUBTRACT WHOLE NUMBERS	14 DECIMAL PROBLEMS (+, -)	15 Story problems (+,-)	16 Commutative, associative & distributive properties	17 MULTIPLICATION (TO 999X99)	18 Multiply decimals	19 CUSTOMARY MEASURE	20 METRIC MEASURE	21 FIND PERIMETER	22 Find area, volume, & surface area	23 Open sentences (x, /)	24 DIVISION (TO 999/99)	25 Division (to 100ths)	26 Open sentences (symbol or number)	27 Open sentences using ( ) & order of operations	28 Estimation (x, /)	29 Story problems (x, /)	30 Prime/composite numbers	31 Identify factors & greatest common factor	32 Points on number line	33 RENAME FRACTION	34 Order/compare fractions & mixed numbers
SHE	DAVID	/	X	/	/	/	/	/	/	/	/	X	X	X	X	X	/	X	/	X	/	X	X	X	X	X	X	/	/	/	X	X	/	/
SHEP	BETH	X	X	X	X	X	X	X	/	X	X	X	X	X	X	X	X	X	/	X	X	X	X	X	X	X	/	X	/	/	X	X	/	X
SILVER	JEROME	X	X	X	X	X	X	X	/	/	X	X	X	X	X	/	X	X	X	X	X	X	X	X	X	X	X	X	/	/	X	X	/	X
ZAMORA	BECKY	X	X	/	/	/	/	X	/	X	X	/	X	X	X	X	X	/	X	X	/	X	/	/	X	X	/	X	/	/	X	X	/	/
% MASTERING OBJECTIVES		90	100	40	80	70	50	70	10	60	80	80	90	90	90	80	70	90	60	80	60	90	50	80	80	90	70	90	30	60	50	90	90	0

Figure 2 (continued)

BEST COPY

Communication Baseline Objective

GRADE 6

oaks rossboro school district

STUDENT PROGRESS REPORT

MARKING CODE	
[X]	MASTERED
[/]	NOT MASTERED
[ ]	NOT INTRODUCED OR NOT TESTED

RAN TU N	WILKINS VIVIAN I	6/3	WILLIAM LIPSCOMB	12/10/82
Student	Teacher	Grade/Section	School	Date Scored

DISD Promotion/Retention (P/R) Standards are printed in "ALL CAPS." TABS-related (exit-level) objectives are marked with an asterisk (•). See Communications Baseline for a detailed explanation of some objectives and complete listing of all objectives.

I. WORD ATTACK SKILLS	FALL	SPR
37.1. Uses possessive s', es' in sentence context.	[/]	[ ]
39. Uses prefixes appropriately in sentence context: super-, anti-, semi-.	[X]	[ ]
40. Uses suffixes appropriately in sentence context: -some, -hood, -ish, -en.	[X]	[ ]
II. VOCABULARY SKILLS		
42. USES CONTEXT CLUES TO DETERMINE UNKNOWN WORDS.	[/]	[ ]
43. CHOOSE A WORD THAT WILL FIT IN A GIVEN SENTENCE.	[X]	[ ]
45. RECOGNIZES AND IDENTIFIES MEANINGS OF SYNONYMS AND ANTONYMS.	[X]	[ ]
III. COMPREHENSION		
Reads with understanding		
I-1. RECOGNIZES SPECIFIC FACTS (DETAILS) STATED.	[X]	[ ]
I-2. RECOGNIZES SEQUENCE OF EVENTS	[X]	[ ]
I-3. •FOLLOWS WRITTEN DIRECTIONS.	[X]	[ ]
I-4. •IDENTIFIES STATED MAIN IDEA.	[X]	[ ]
II-1. •FORMS CONCLUSIONS.	[X]	[ ]
II-2. MAKES COMPARISONS.	[/]	[ ]
II-3. IDENTIFIES IMPLIED MAIN IDEA.	[X]	[ ]
II-5. PREDICTS EVENTS AND OUTCOMES.	[X]	[ ]
II-6. IDENTIFIES CHARACTER FEELINGS.	[X]	[ ]
III-1. •RECOGNIZES FACT AND OPINION.	[X]	[ ]
IV. SPELLING		
49.1. •SPELLS MOST COMMONLY USED WORDS.	[X]	[ ]
60. Adds endings to words which end in y.	[X]	[ ]
V. COMPOSING SKILLS		
106.1. •Writes a paragraph consisting of a topic sentence and relevant details to fit a specific audience and purpose.	[ ]	[ ]
99. •IDENTIFIES IRRELEVANT SENTENCES IN A PARAGRAPH.	[X]	[ ]
104. •Identifies and corrects sentence fragments.	[X]	[ ]
82. CAPITALIZES THE NAME OF CHURCHES AND SYNAGOGUES, ORGANIZATIONS AND COMPANIES.	[X]	[ ]

V. COMPOSING SKILLS (continued)	FALL	SPR
91. •USES COMMAS TO SEPARATE PHRASES IN A SERIES.	[X]	[ ]
90.1. Uses exclamation marks after sentences, words, and phrases of exclamation.	[X]	[ ]
VI. USAGE		
110. Uses pronouns correctly.		
- Selects the correct pronouns for use in compound subjects and objects. e.g., Jimmy and (I, me) are neighbors. Look at (I, me).	[X]	[ ]
- Selects a pronoun to agree with its antecedent in a given sentence (including indefinite pronouns).	[X]	[ ]
109. Uses verbs correctly.		
•- Selects the correct form (person and number) of a verb that agrees with the subject (including compound subjects). e.g., the puppy and its mother (eat, eats) too much.	[/]	[ ]
•- SELECTS THE CORRECT VERB TO USE IN A SENTENCE (LET, LEAVE, TEACH, LEARN, RISE, RAISE, SIT, SET).	[X]	[ ]
•- SELECTS THE CORRECT FORM (PAST, PAST PARTICIPLE) OF BOTH REGULAR AND IRREGULAR VERBS. E.G., I HAVE (DID, DONE) THE WORK.	[X]	[ ]
111.1. Uses adjectives correctly.		
- CHOOSES BETWEEN -ER AND -EST FORM OF AN ADJECTIVE WHEN PRESENTED WITH BOTH FORMS IN CONTEXT.	[/]	[ ]
- Uses these and those (as adjectives) correctly in sentences. e.g., Give me some of (those, them) cookies.	[X]	[ ]
112. •CORRECTS DOUBLE NEGATIVES CONTAINING THE WORDS NO AND NOT.	[X]	[ ]
VII. REFERENCE		
121. USES A DICTIONARY TO CORRECT MISSPELLED WORDS AND TO DEFINE A WORD USED IN CONTEXT.	[X]	[ ]

Fall Conference Date: _____

Spring Conference Date: _____

_____
Parent/Guardian

_____
Parent/Guardian

Figure 3. Sample student progress report for communications.

BEST COPY

CLASS PROFILE

COMMUNICATIONS 6

School: WILLIAM LIPSCOMB

Teacher: WILKINS VIVIAN I

STUDENT NAMES		37.1 Uses plural possessive	38 Uses prefixes	40 Uses suffixes	42 USES CONTEXT CLUES	43 CHOOSES WORD IN CONTEXT	45 SYNONYMS & ANTONYMS	I-1 SPECIFIC FACTS	I-2 SEQUENCE OF EVENTS	I-3 = WRITTEN DIRECTIONS	I-4 = STATED MAIN IDEA	II-1 = FORMS CONCLUSIONS	II-2 MAKES COMPARISONS	II-3 IMPLIED MAIN IDEA	II-5 PREDICTS EVENTS/OUTCOMES	II-6 CHARACTER FEELINGS	III-1 = FACT & OPINION	49.1 = SPELLS COMMON WORDS	80 Adds endings to words	106.b1 Writes for purpose & audience	99 = IRRELEVANT SENTENCES	104 = Sentence fragments	82 CAPITALIZES NAMES	91 = USES COMMAS	90.1 Uses exclamation marks	110 Selects plural pronouns	110 Selects pronoun in context	109 = Selects verb form	108 = SELECTS VERBS IN CONTEX	109 = SELECT VERB TENSE	111 = CHOOSES -ER, -EST /COMPARATIVE	111.1 Uses these & those	112 = CORRECTS DOUBLE NEGATIVE	12.1 USES DICTIONARY		
RAN	TU N	/	X	X	/	X	X	X	X	X	X	X	/	X	X	X	X	X	/	X	X	X	X	X	X	/	X	X	/	X	X	X	X		
BLEV	DARREL V	/	X	X	X	X	/	X	X	X	X	X	X	X	X	X	X	X	/	X	X	X	X	X	X	X	X	X	/	X	X	X	X		
DELO	SARA	/	/	X	X	/	/	X	/	/	X	X	X	X	X	X	/	X	/	X	/	X	X	X	X	/	X	X	/	X	X	X	X		
EIRII	KATHREN	/	/	X	X	X	X	X	X	X	X	X	X	X	X	/	X	X	X	X	X	X	X	X	X	X	X	X	/	X	X	X	X		
GANN	CHARLESE	X	X	X	X	X	X	X	/	/	X	X	X	X	X	X	X	X	X	X	X	X	X	X	/	X	X	X	/	X	X	X	/		
GONZA		/	X	X	X	X	X	X	X	X	X	/	X	X	/	X	X	X	X	X	X	X	X	X	X	X	X	X	/	X	X	X	X		
HASTI	ROBERT A	/	X	X	X	X	/	X	X	/	X	X	X	X	/	/	X	X	X	X	X	/	/	X	/	X	X	X	/	X	X	X	/		
HINE	TRAVIS S	/	X	X	X	X	/	X	/	X	X	/	X	X	X	/	X	X	X	X	X	X	X	X	/	X	X	X	/	X	X	X	X		
KRAEG	AMY	/	/	/	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	/	X	X	X	/	X	X	X	X		
LOCAN	RAUL A	/	X	X	X	X	X	X	X	X	X	X	X	X	/	X	X	X	X	X	X	X	X	/	/	X	/	/	/	X	X	X	X		
MIN	STACE	X	X	X	X	X	X	X	X	X	X	/	X	X	/	X	/	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
RIO	ELECIA	/	X	X	X	X	/	X	X	X	X	X	X	X	/	X	X	X	X	X	X	X	X	/	X	X	X	X	X	X	X	X	X		
ROSA	JAMES A	/	/	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	/	X	X	X	X	X	X	X	X	X	X	X		
VILLAL		/	/	X	X	X	/	X	/	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	/	X	X	X	X	X	X	X			
% MASTERING OBJECTIVES		14	64	93	93	93	57	100	71	79	100	79	93	100	79	64	86	100	86	0	100	93	86	93	86	50	100	86	100	86	21	93	93	71	

Figure 4. Sample class profile for communications.

BEST COPY

MINUTIA AND TRIVIA OF DEVELOPMENT AND IMPLEMENTATION

Developing and implementing a learning system like the one just described, required the solution of numerous unseen and unconsidered problems. When the problems to be outlined and discussed are considered individually and out-of-context they seem quite trivial. However, if they are not solved they can hinder or cause the entire development and implementation process to halt. The following discussion will outline some of the trivial day-to-day activities and problems that we have experienced related to development of an assessment system.

Test development tasks seem to proliferate minutia and trivia. Thorndike (1971) addresses some of the trivia related to reproducing a test. However, for the most part, these trivia go unmentioned in measurement courses and professional meetings. Consequently, novice and would be test developers have no real idea of what is in-store for them. We feel it is our duty to forewarn and inform these people about some potential problems and possible solutions, or dissuade them from becoming test developers.

There are two areas in which trivia tend to proliferate that we would like to discuss: (a) item formating and (b) assembling a copy of the test.

Item formating. To preclude or avert problems related to item formating, standards should be set before your writers and typists begin. Standards which should be considered include the following:

1. Margins
2. Size of type at each level (e.g., large type for primary grades, smaller type for upper grades)

3. Use line drawings (because full figures do not reproduce well).

4. Type of items allowed (e.g., multiple-choice, matching, true-false)

5. Arrangement of items (e.g., two column vs. strung-out arrangement)

6. Arrangement of choices.

A	C	A	B	A	
B	D	C	D	B	[A B C D]
				C	
				D	

7. Labeling of choices. Numbers or letters; if letters, then capitals or lower case. At some grade levels, should numbers be used with language arts, while letters are used with mathematics.

8. Punctuation used with choice label. Examples: A., A), (A), A-

9. Labeling of choices in a Spanish edition.
Examples: (A, B, C, D) (A, B, C, CH) (A, B, C, Ch)

10. Selecting size of oval for primary grades.

Test assembly. To facilitate assembling the test copy, a number of considerations related to graphics and copying need to be made. In many cases early decisions will promote orderly planning, development, and implementation; thus averting potential crises. Specific considerations in assembling the test copy include the following:

1. Attaching graphics
   a. Glue sticks
   b. Tape
   c. Spray
   d. Rubber Cement
2. Correcting errors
   a. Word processor
   b. Liquid paper (one bottle for originals, one bottle for copies, one bottle for pen and ink)

3. Finding or drawing pictures, diagrams, or figures
    a. Templates
    b. Non-copyrighted material (Clip Art)
    c. Using artists
    d. Line drawings
4. Copying machines and features
    a. Paste-up suppression
    b. Reduction
    c. Multiple reduction
5. Assembling test pages from individual items
    a. Use of document carriers with hard copy
    b. Use of wordprocessing (move, copy, supercopy)
6. Designing scannable booklets and answer sheets.

As a concluding note, these lists are not intended to be all inclusive of the details and tasks related to test development, but an attempt to share some of the things we have learned from our experiences. Our hope is that novice test developers might learn from our experiences.

Conclusion

In summary, we feel that the DISD assessment system "works." Information resulting from the system is used by personnel at all levels from the classroom teacher to the general superintendent. Even greater support for the contention that the assessment system works is the wide, voluntary use of the classroom tests by teachers and schools. During the 1982-83 school year over 65,000 tests were scored at grade levels 4-6. As of January 1984, 59,148 tests had been scored during the 1983-84

school year. The student population in grades 4-6 is approximately 30,000. Thus, the classroom tests are being used periodically by many teachers to measure individual student progress and to evaluate their instructional strategies and processes. This wide use also illustrates the importance of linking the assessment system directly to an overall system of instruction and school improvement planning. When schools are held accountable for producing increases in student achievement, they begin to see the necessity of monitoring progress on an interim basis. They also begin to perceive the necessity of building their program of instruction on a foundation of accurate information regarding the skill levels of the students for whom they are planning.

The strongest component of the DISD's total assessment program as it relates to instruction is the classroom test component. There are at least six reasons for this:

1. The tests have considerable credibility since they match the curriculum exactly.

2. The tests can be used at any time to assess the impact of the school improvement plan instructional strategies on priority skills.

3. The tests can be tailored specifically to the content and sequence of each teacher's instruction.

4. Care has been taken so that computer-generated reports are in the same format as was previously used by the teachers, thus making an easy transition.

5. The tests and computer services are convenient to use and lead to a reduction of teacher paperwork.

6. While not designed strictly for accountability purposes, these tests provide information which can give a preliminary measure of accountability for specific targeted areas so that adjustments can occur as necessary.

APPENDIX A

Test Development Activities for Survey and Classroom Tests

The Dallas Independent School District (DISD) has been working through a five-year instructional improvement process. One aspect of this process entailed establishing minimum standards for student learning at each grade level. These standards (also referred to as promotion/retention objectives) are a subset of the Baseline Curriculum. Attainment of mastery of these objectives is required. A student's performance on these objectives is used as a guideline in determining whether he or she will be promoted or retained.

A second aspect of the instructional improvement process entailed developing a criterion-referenced testing system. Implementing an objectives-based curriculum like the Baseline involves placing students at the appropriate objective level, instructing the students, and assessing post-instruction performance. After instruction, teachers must decide whether to provide alternative activities or enrichment exercises, or to continue in the sequence of objectives. To make decisions concerning placement, mode of instruction, and progress, teachers need a variety of evaluation tools including carefully crafted, criterion-referenced tests. These tests consist of items that are matched to specific objectives. Hence, scores on the tests can be used by teachers to determine how well or how poorly their students are performing on the objectives of interest.

Members of the Test Development staff of the Department of Research, Evaluation, and Information Systems have developed the testing component

of the instructional improvement process. The testing component, Assessment of the Baseline Curriculum (ABC), consists of two major products. These are (a) survey tests that will be used to provide standardized measurement of students' performance on the promotion/retention objectives and (b) sets of classroom tests and an attendant scoring system that can provide information to teachers for making day-to-day instructional decisions.

This appendix will describe both of these testing components. It will outline the development of the survey tests. It will also present the organization, implementation, and utility of the classroom tests.

SURVEY TESTS

During the 1980-81 and 1981-82 school years, the ABC Survey Tests were developed for grades 1-6. These tests measure the promotion/retention objectives at each level of the Baseline Curriculum.

All of these Survey Tests were administered District-wide at one time during the school year, and students received the level of the test that corresponded to their grade level. That is, students were tested on the objectives that are the standards for their grade.

This section of the paper describes the development of the ABC Survey Tests. For organizational purposes, the process has been divided into seven broad categories. These are: Objectives, Domain Specifications, Test Items, Cut Scores, Reliability, Validity, and Test Score Reporting System.

Objectives

The test construction process began during the spring of 1980. A group of Early Childhood teachers and instructional specialists was

convened by the Curriculum Development Department. This group selected promotion/retention objectives in mathematics and language arts for Levels 1-3. During the spring of 1981 another group of teachers selected promotion/retention objectives in mathematics and language arts for Levels 4-6.

Domain Specifications

Using the lists of promotion/retention objectives as starting points, members of the Test Development group worked with classroom teachers, instructional specialists, and extra-District consultants to draft, review, and finalize detailed specifications of what test item content should be and how test items should look. These detailed specifications are referred to as domain specifications. Figure 1 presents an illustration of a domain specification for a mathematics objective.

During June of 1980, Test Development personnel trained classroom teachers to write domain specifications for grades 1-3 during June of 1980 and for grades 4-6 during June of 1981. The teachers then proceeded to construct initial drafts of the domain specifications for the promotion/retention objectives in mathematics and language arts.

During the 1980-81 and 1981-82 school years, the domain specifications were reviewed by (a) Test Development personnel, (b) Learning Division personnel, (c) Curriculum Development personnel, and (d) measurement specialists from the School of Education at the University of Massachusetts. The suggestions from these reviews were compiled, analyzed, and when constructive, incorporated into the domain specifications.

BEST COPY

Content Area: Mathematics
Level: 2
Objective Number: 69
Domain Specification Number: 1 of 1

- General Description

Find the value of a collection of money less than $1.00.

Model Test Item and Directions

Directions (to be read to the student)

Find the sailboat. Look at the coins next to the sailboat. Fill in the oval under the amount of money shown by the coins.

Test Item (to appear in the student's test booklet)

45¢ 46¢ 55¢ 60¢

Figure 1. Domain Specification for a Mathematics Objective.

BEST COPY

Eligible Test Content

1. The coins will be picture copies of the actual coins.

2. The coins will include half-dollars, quarters, dimes, nickels, and pennies. The half-dollars will be Kennedy or Franklin, the quarters will be Washington, the dimes will be Franklin Roosevelt, the nickels will be Jefferson, and the pennies will be Lincoln.

3. The head or the tail side of the coins may be visible. If possible, the writing on the coins will appear right-side-up relative to the bottom of the page.

4. The collection of coins will range from two to five. Their total value will be less than $1.00.

5. There will be at least two denominations of coins in the collection.

6. The collection of coins does not necessarily have to appear in row form. Also, the coins can appear in any order. However, the coins will not touch.

Eligible Response Content and/or Scoring Rules

1. There will be four response choices, only one of which is correct.

2. The response choices will be arranged horizontally in ascending order.

3. The response choices will all use the "cents" (¢) sign and not "dollar sign decimal point" ($.). They will range from 5¢ to 99¢.

4. The incorrect choices should represent the most frequent student errors. For example, one of the incorrect choices could be a cents amount which results from the improper identification of coins (e.g., identifying a nickel as a quarter or vice versa).

Specification Supplement

1. This objective is similar to, but less difficult than, TABS Objective 9 at Grade 3.

2. It is recommended that five items (each involving a different collection of coins) be used to measure this objective. Mastery status should be defined as four out of the five items correct.

Figure 1. (Continued) Domain Specification for a Mathematics Objective.

The Curriculum Development Department pilot tested the promotion/retention objectives for grades 1-3 in nine Early Childhood Centers during the 1980-81 school year. The pilot-testing for grades 4-6 took place in ten Intermediate schools during the 1981-82 school year. The results of this piloting led to several slight modifications in the promotion/retention objectives. The final step in the preparation of the domain specifications was to incorporate these minor modifications.

Test Items

The next step in the test construction process was the development and the refinement of test items. Using the domain specifications as blueprints, members of the Test Development group wrote test items to measure each promotion/retention objective at Levels 1-6. After the test items were written, they were reviewed by Test Development personnel. The review process focused on determining (a) the extent to which the test items "matched" the specification, and (b) the technical quality of the test items. The results of the item review process were compiled and analyzed, and the items were modified in accordance with the constructive suggestions.

As was mentioned earlier, nine schools were involved in the pilot testing of the promotion/retention objectives. These schools also participated in the tryout of the Survey Tests during the springs of 1981 and 1982. The ABC Survey Tests were administered by Research and Evaluation personnel. Most of the items in these tests were multiple choice. Several were of the constructed-response type. Students at grades 1 and 2 responded in their test booklets, while third through

sixth graders used an answer sheet to record their multiple-choice answers. Finally, all of the tests were untimed.

During the tryout of the Survey Tests, the classroom teachers reviewed the test items. The teachers answered questions about the technical quality, appropriateness, and difficulty of each test item.

The information from the tryout of the tests and the teacher review of the test items was used to refine the Survey Tests. The item review and selection procedure focused on identifying the items measuring each objective that had about the same difficulty levels. Test items that had substantially different difficulty levels from the majority of items measuring an objective were identified, reviewed, and, where possible, modified. In addition, the comments made by the teachers who reviewed the test items were incorporated in the item selection procedure. The majority of these comments resulted in the modification of test items. When the items were not able to be modified, other items were developed from the specifications, and these items were substituted for the original items.

Cut Scores

Cut scores are standards that help teachers interpret test scores. Typically, students who score at or above a particular cut score are classified as masters, and students who score below the cut score are classified as nonmasters. Cut scores are set by using a logical and reasonable method.

Actually, two methods were used to set cut scores. The first method required teachers to estimate the percentage of minimally promotable students who could answer each item correctly. Essentially, the cut

score for a set of items measuring an objective was taken as the sum of the estimated percentage for each item. Note that this technique for collecting data to set cut scores was similar to that originally suggested by Angoff (1971).

The second method required the teachers to classify each of their students into two groups (i.e., master of the objective vs. nonmaster of the objective; promotable vs. not promotable). The cut score was logically taken as the score between the average scores of the two groups of students. Note that this technique was similar to the contrasting groups method developed by Zieky and Livingston (1977).

In most instances the two techniques resulted in the same cut score. When the techniques did not agree, the students were classified by both cut scores. Each classification was then compared to the teacher classification. The cut score which resulted in the highest level of agreement with the teacher classification was selected as the cut score to be used in the test.

Reliability

The approach to reliability that was adopted was consistent with how the scores from the Survey Tests were to be used. Scores on each objective as well as the total test score would be compared against the cut scores, and mastery and nonmastery decisions would be made. These decisions would be used by administrators and teachers as indicators of students' status on the promotion/retention objectives. The reliability that is of concern is the consistency of the mastery decisions over repeated testings. That is, if the Survey Test was administered twice without intervening instruction and the same standards of performance

were used, would similar decisions be made? Would the student who was classified as a master on one test administration also be classified as a master on the second test administration? Or would the nonmaster on administration one also be a nonmaster on administration two?

To assess decision consistency, the scores from one test administration were used to simulate the scores on the second test administration. Both sets of scores were compared against the cut score, and the consistency of the mastery decisions over the two administrations was estimated.

The results of this analysis indicated that the mastery decisions based on the test scores from the sets of items developed to measure the promotion/retention objectives were consistent. For example, the average estimated consistency of mastery decisions over two test administrations was 88% in mathematics and 86% in language arts (Shsehan, Melican, & Arrasmith, 1981). Thus, one can have confidence that the mastery decisions made as a result of one test administration would, in all likelihood, be made in a similar manner as a result of a second test administration.

Validity

Two types of validity are important for the Survey Tests. These are content validity and construct validity. Content validity refers to the match between the test items and the objectives they were designed to measure. Construct validity refers to the degree to which explanatory concepts account for scores on the tests. That is, do the test scores accomplish their intended uses? Evidence relating to each of these types of validity was compiled throughout the test construction process.

Content Validity

The entire test construction process helped ensure the content validity of the Survey Tests. This process was concerned with the validity of test items, their technical quality, and their representiveness.

Item validity. Domain specifications were drafted, reviewed, and finalized to precisely define the content covered by the promotion/retention objectives. Using the specifications as blueprints, test items were written to measure each objective. Test Development personnel then judged the extent to which each test item measured the objective for which it was written. When the items did not perfectly match the specifications, they were modified so that they would match.

Pilot-test data were also used to create content-valid test items. Item difficulty values within a set of items measuring a particular objective were analyzed for aberrant values, and, hence, aberrant items. These items were examined for flaws and, if found, corrected or replaced by other items developed from the specifications. It is important to note that the empirical data were used for the correction of items and not for their elimination. Thus, the approach to item validity used both judgmental and empirical methods to ensure that the content of the test items matched the objectives they were designed to measure.

Technical quality. Part of the item review process consisted of Test Development personnel analyzing the test items for their technical quality. When items were found that violated sound item-writing rules, the defects were corrected.

Representativeness. An important aspect of content validity is the representativeness of the set of items designed to measure the objective

(i.e., the item pool). That is, as a group, do the items in the pool measure all aspects of the content domain? Item representativeness was ensured by stating in the domain specification the number of items that had to be written to cover each important aspect of the domain. For example, if addition with regrouping and addition without regrouping were important facets of a particular domain, the specification would state the number of items both with and without regrouping that had to be written.

This is an extremely important, yet often overlooked, aspect of content validity. Taken individually, each item in a pool may match the objective for which it was written, yet, as a group, the pool of items may not measure an important part of the objective. Thus, the specification must list the number of items of each important type that have to be written.

Construct Validity

The American Psychological Association (1966) states that construct validity is evaluated by investigating what psychological qualities a test measures; i.e., by determining the degree to which certain explanatory concepts or constructs account for performance on the test. Given that the Survey Tests were designed to measure the constructs defined by the promotion/retention objectives, two types of data that are relevant to their construct validity were accumulated. The first is their decision accuracy or the extent to which test decisions are consistent with the decisions from another method of measurement. The second is a broad class of necessary, but not necessarily sufficient, conditions for construct validity.

Decision validity. The Survey Tests were designed to measure whether students are able to master each promotion/retention objective or each "essential" Kindergarten objective. In view of this use of the tests, it was important to know the extent to which the test mastery-nonmastery decisions reflected actual classroom performance. The measure of classroom performance that was used was teacher judgment as to whether or not each student was a master of a particular objective. Teacher mastery/nonmastery decisions on specific objectives were extracted from student progress forms. The agreement between these teacher decisions and the test mastery/nonmastery decisions was then determined.

The mastery decisions made from the results of the Survey Tests agreed with the teacher decisions more than 70% of the time. Specifically, the average percentage of agreement was 77% for the mathematics tests and 79% for the language arts tests (Sheehan et al., 1981). Given that factors other than test performance (e.g., classwork, homework, etc.) can influence the teachers' mastery decisions, the classification accuracy of the tests seems to be satisfactory.

Necessary conditions for construct validity. The Survey Tests were designed to measure student performance on the District's promotion/retention objectives at gardes 1 to 6. These objectives represent relatively basic or minimum competencies. As such, three necessary, but not sufficient, conditions for their construct validity would be (a) a high degree of internal consistency, (b) a low level of difficulty, and (c) a high completion rate.

In actuality, all three conditions were satisfied. Internal consistency coefficients indicated that the Survey Tests tended to be very homogeneous. All of the items within a test appeared to be

measuring the same construct. Performance levels were very high. That is, as expected, the tests were not very difficult. Finally, almost all of the students completed the tests. Completion rates exceeded 90%.

CLASSROOM TESTS

During the 1981-82 school year, two major products were produced in the area of classroom tests. These were (a) ABC Classroom Tests for Levels 1-6 (English editions) and (b) ABC Classroom Tests for Levels 1-3 (Spanish editions). In addition, the English versions of the ABC Classroom Tests for Levels 1-3 which were developed during the 1980-81 school year were modified to incorporate Baseline revisions and were reorganized into booklet format. All of these tests measure student attainment of objectives in DISD's Mathematics and Communications Baseline Curricula. They also provide a standard District-wide measure of student progress. Because most of the tests measure only one objective, teachers can select the tests of the Baseline Objectives in the order most appropriate for their instruction.

Organization and General Implementation

ABC test materials consist of student's booklets and teacher's manuals for each level of the curriculum and for each content area. The student's booklets assess all objectives that appear on the Student Progress Report. (The Student Progress Report is a form that is used in parent conferences.) A table of contents is provided to help locate tests which measure specific objectives. Tests for the Promotion/Retention Objectives are included in each booklet and are identified by labels (P/R) in the table of contents and in the tests.

The teacher's manuals contain a test management system, a test answer key, and a copy of the student's booklet. The test management system describes the purpose of the tests, their organization, several classroom applications, their implementation, and a scoring system. The test answer key includes a recommended minimum passing score.

The student's booklets are consumable at Levels 1-3. At Levels 4-6 the booklets are nonconsumable. The students respond on separate answer sheets. Any of these materials can be ordered free of charge by individual schools from the School Service Center.

Applications of the Classroom Tests

Generally there are four applications of the ABC Classroom Tests. These are: (a) student placement, (b) specific assessment of student learning, (c) survey testing, and (d) development of testing skills. A testing situation can involve more than one of these applications; however, for clarity the four applications will be discussed separately.

Student Placement

In the first stage of an instructional cycle, the placement of the student in the appropriate level of instruction is important. While this placement should not depend entirely on the results of tests, these scores can be helpful. The ABC Classroom Tests offer two strategies for placement of students: (a) existing progress charts and previous test results, and (b) selective placement tests.

To avoid unnecessary testing, placement should be based whenever possible on existing information from progress charts and previous testing (e.g., MRT and ITBS). The existing information will often be

sufficient for the placement of students in the subsequent instructional sequences.

If, however, previous student information is not available, the ABC Classroom Tests can be used for student placement by selective testing. The general strategy for such testing is to focus first on a few "culminating" objectives in order to place students in a grade level and then focus on the remaining objectives to identify specific instructional needs. The Test Management System outlines a suggested procedure for this selective testing in the Mathematics, Reading, and Writing Curricula.

Assessment of Student Learning

The ABC Classroom Tests are designed primarily for assessing student attainment of the Mathematics and Communications Baseline Objectives. The tests can be used as pretests prior to instruction and/or as posttests following instruction. Pretest information will be useful in planning instruction. Posttest information should either (a) confirm that the student has attained mastery and should proceed to the next objective or (b) help direct remedial instruction.

It should be noted that the pre/post testing is only a suggestion. The teachers and instructional leaders should adapt the Classroom Tests to their own instructional methods. The tests have been designed to have this flexibility.

Survey Testing

The ABC Classroom Tests can be grouped to form survey tests covering a set of objectives. An important instance of this application is the ABC Survey Test covering promotion/retention objectives. A parallel form of this District-wide Survey Test can be constructed from the Classroom

Tests. The test items which are similar to those in the District-wide Survey Test have been designated by the label P/R in the Table of Contents and on the appropriate tests. These test items can be combined to produce a test which will be very similar to the District-wide ABC Survey Test.

In addition, the teachers or instructional leaders may want to design survey tests covering previously instructed objectives. By selecting the appropriate test items, such tests can be produced.

Development of Testing Skills

In order to help students become familiar with the types of test items used in the ABC Classroom Tests and to develop good test-taking skills, a practice test has been included with the Level 1-3 tests. This practice test should be given to students prior to the first ABC Classroom Test and, if needed, prior to any subsequent Classroom Tests. The practice tests are important for two reasons: (a) students will be taught how to show accurately what they can do on the ABC Classroom Tests, and (b) students will be taught important study skills that will be useful in solving problems and in taking other tests.

With the use of the practice tests, the students will become accustomed to answering the type of questions that are contained in the ABC Classroom Tests. The ABC Tests contain a variety of testing formats and test-related language. Thus, by using the ABC Tests, students will become proficient in using a variety of test formats and test-related language. By using both the ABC Practice Tests and the Objective Tests, students will learn how to read and answer the questions, how to mark their answers, and how to do their best on other tests.

Implementation of the Classroom Tests

The ABC Classroom Tests have been designed to be administered to a group of students. This does not exclude the use of the tests with an individual student. In fact, in some instances, individual administration of the tests may better help the teacher assess student learning. In addition, the ABC test items could be adapted to writing the test on a chalkboard. The teacher using this strategy may want to leave the test items open-ended and not write the optional responses. Furthermore, some of the test items may be presented orally to the students. Again, the test items may be left open-ended. The point to be made is that the teacher should feel free to adapt and use the ABC Tests as best fits his or her instructional needs.

REFERENCES

American Psychological Association. Standards for educational and psychological tests and manuals. Washington: APA, 1966.

Angoff, W. H. Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), Educational measurement. Washington, D.C.: American Council on Education, 1971.

Sheehan, D. S., Melican, G., & Arrasmith, D. G. Technical Manual for the ABC (Assessment of Baseline Curriculum) Survey Tests - Levels 1-3. Dallas, Texas: Dallas Independent School District, Department of Research and Evaluation, 1981.

Zieky, M. J., and Livingston, S. A. Manual for setting standards on the Basic Skills Achievement Tests. Princeton, N.J.: Educational Testing Service, 1977.

Thorndike, R. L. Reproducting the Test. In R. L. Thorndike (Ed.), Educational measurement. Washignton, D.C.: American Council on Education, 1971.